

Table 4.10 **Comparison of Moments in Different Units of Measurement**

Moment	Height (in.)	Height (ft)
m'_1	65.37 in.	5.450 ft
m_2	11.24 in. ²	0.078 ft ²
m_3	-3.83 in. ³	-0.002 ft ³
m_4	283.20 in. ⁴	0.014 ft ⁴

If we raise the differences between the x_i and \bar{x} to a high power, the “big” numbers will have a much bigger effect on the sum of the powers than the more numerous observations in the middle that produce small differences. When we calculate m_4 , and if our intuition has been reliable, we should see the values of m_4 increase as we examine the distributions in the order we specified.

The results of these calculations are in Table 4.9. Unfortunately, the results do not match our expectations. The fourth moments of the distributions that we thought would be the smallest are not so, nor are the fourth moments we thought would be the largest so. We need to go back to the drawing board.

4.6 Standardized Moments

You may have noticed another problem with the third and fourth moments: what’s large? It is not very helpful to say that m_4 is large for highly peaked distributions if we do not know what constitutes “large.” Our cavalier attitude needs reassessment.

Even for m_3 we have a problem. If we want to compare two histograms that are both skewed to the left, for example, can we unambiguously say that one of them is more skewed than the other on the basis of the values that we obtained for m_3 ? Yet another problem is raised if we try the next experiment.

In Table 4.10, we recorded the first four moments on the heights of enrollees in a fitness class measured in inches. We converted these numbers into feet. We are essentially dealing with the same histogram of heights, so our conversion to feet was merely a re-scaling of the data; we should conclude that we have the same shape of histogram.

The mean is easy enough to understand; m'_1 in feet is just one-twelfth of m'_1 in inches. But the second moment is not so easy to interpret, and except for signs, m_3 and m_4 appear to be completely different between the two measurements. This is not desirable at all; our results should not depend on our choice of the units of measurement if we are to create a generally useful measure of shape.

Now suppose that we measure heights not from zero as we have done so far, but from 60 inches; the idea is that we are interested mainly in the deviations of actual heights from 5 feet. If we now recalculate all our moments, we get

$$\begin{aligned}
 m'_1 &= 5.37 \text{ in.} \\
 m_2 &= 11.24 \text{ in.}^2 \\
 m_3 &= -3.83 \text{ in.}^3 \\
 m_4 &= 283.20 \text{ in.}^4
 \end{aligned}$$

Comparing the means we get what we might have expected: m'_1 for the data after subtracting 60 inches is just our original mean value of 65.37 inches less 60 inches. But what may be surprising is that m_2, m_3, m_4 are all exactly the same as the moments we obtained from the original data!

A moment's reflection may help you to see why this is so. Look at the definition of m_2 , for example:

$$m_2 = \frac{\sum (x_i - \bar{x})^2}{N}$$

Define a new variable y_i by

$$y_i = x_i - 60$$

because we subtracted 60 from the height data in inches.

To emphasize the variable whose moment is being taken, let us change notation slightly to $m_r(y)$, $m_r(x)$, or $m_r(w)$ to represent the r th moment about the mean for the variables y , x , and w respectively.

Now calculate the mean for the new variable y :

$$\begin{aligned} m'_1(y) &= m'_1(x_i - 60) \\ &= m'_1(x_i) - m'_1(60) \\ &= 65.37 - 60 = 5.37 \end{aligned}$$

If you have any trouble following these steps, a quick look at Appendix A will soon solve your difficulty.

Now we calculate $m_2(y)$.

$$\begin{aligned} m_2(y) &= \frac{\sum (y_i - \bar{y})^2}{N} \\ &= \frac{\sum [(x_i - 60) - (m'_1(x) - 60)]^2}{N} \end{aligned}$$

where we have substituted

$$x_i - 60 \text{ for } y_i; \quad m'_1(x) - 60 \text{ for } m'_1(y)$$

but,

$$\begin{aligned} &\frac{\sum [(x_i - 60) - (m'_1(x) - 60)]^2}{N} \\ &= \frac{\sum (x_i - \bar{x})^2}{N} \end{aligned}$$

because the two "60s" cancel in the previous line: $-60 - (-60) = -60 + 60 = 0$. We have shown that the second moments for the variables x and y , where y is given by $y = x - 60$, are the same.

This result will hold for all our moments about the mean. This property is called *invariance*; the moments $m_2, m_3, m_4 \dots$ are all said to be **invariant** to changes in the

origin. This means that you can add or subtract any value whatsoever from the variable and the calculated value of all the moments about the mean will be unchanged.

The qualification “about the mean” is crucial; this is not true for the moments about the origin.

We have solved one problem; we now know that, beyond the first moment, we need to look at moments that are about the mean and so invariant to changes in the origin. But the problem of interpreting the moments, detected with the heights, goes further than a change of origin. When we changed from inches to feet, we did not change the origin, it is still zero, but we did change the scale of the variable. But “scale” is like spread; a bigger scale will produce a larger spread and consequently a larger value for m_2 , our measure of spread. Further, our measure of spread is squared. Reconsider the expression for m_2 :

$$m_2(x) = \frac{\sum (x_i - \bar{x})^2}{N}$$

and consider changing from measuring x_i in inches to a variable y_i that is height measured in feet. We do this by dividing the x_i entries by 12:

$$\begin{aligned} \frac{\sum (y_i - \bar{y})^2}{N} &= \frac{\sum \left(\frac{x_i}{12} - \frac{\bar{x}}{12}\right)^2}{N} \\ &= \frac{\left[\frac{\sum (x_i - \bar{x})^2}{N}\right]}{144} \end{aligned}$$

where $144 = 12^2$. Or more generally, if $y_i = b \times x_i$, for any constant b , then

$$m_2(y) = b^2 \times m_2(x)$$

Even more generally, if $y_i = a + (b \times x_i)$, for any constants a and b ,

$$m_2(y) = b^2 \times m_2(x)$$

If we now apply this same approach to the higher moments, we will discover that, whenever

$$\begin{aligned} y_i &= a + (b \times x_i) \\ m_2(y) &= b^2 \times m_2(x) \\ m_3(y) &= b^3 \times m_3(x) \\ m_4(y) &= b^4 \times m_4(x) \end{aligned}$$

and so on.

This business of changing variables by adding constants and multiplying by constants may still seem a little mysterious, but a familiar example will help. You know that temperature can be measured either in degrees Fahrenheit or in degrees Celsius. The temperature of your sick sister is the same whatever units of measurement you use; that is, how much fever your sister has is a given, but how you measure that temperature, or how you record it, does depend on your choice of measuring instrument.

To reexpress the idea, your choice of the units of measurement alters the measurement, but it clearly does not alter the degree of fever.

One choice is to measure in degrees Fahrenheit. Suppose that you observe a measure of 102 degrees Fahrenheit. That observation is equivalent to a measured temperature of 38.9 degrees Celsius. As you may remember from your high school physics, degrees Fahrenheit are related to degrees Celsius by

$$\text{deg.F} = 32 \text{ deg.F} + \left(\frac{9}{5}\right) \times \text{deg.C}$$

but this is just like

$$\text{deg.F} = a + b \times \text{deg.C}$$

where $a = 32 \text{ deg.F}$ and $b = \frac{9}{5}$.

The origin for degrees Fahrenheit, relative to degrees Celsius, is 32°F and the scale adjustment is to multiply by $\frac{9}{5}$. So, if you were interested in the shape of histograms of temperatures, you would not want your results to depend on how you measured the data [except, of course, for measures of location (that depend directly on origin and scale) and measures of scale or spread that should depend only on scale, not on the choice of origin.]

We can reexpress our results so far by saying that m'_1 indicates the chosen origin for the units of measurement and that $\sqrt{m_2}$ indicates the chosen scale of measurement; “ $\sqrt{\quad}$ ” is the traditional square root sign and means take the square root of its argument.

We now have the answer to our problem of trying to decide whether a third or a fourth moment is large or small. We also have our answer to the question of how to ensure that our measures of shape do not depend on the way in which we have measured the data. If we divide the third and fourth moments about the mean by the appropriate power of m_2 , we will have a set of moments that will be invariant to changes in scale. This is equivalent to picking an arbitrary value for b in the equation $y_i = a + (b \times x_i)$.

We conclude that to discuss shape beyond location and scale in unambiguous terms, one has to measure shape in terms of expressions that are invariant to changes in origin or in scale. As we saw, any change in scale in the variable x is raised to the power 3 in the third moment and to the power 4 in the fourth moment. Thus an easy way to overcome the effects of scale is to divide m_3 by $(m_2)^{3/2}$ and m_4 by $(m_2)^2$.

Define $\hat{\alpha}_1$ and $\hat{\alpha}_2$, the **standardized** third and fourth **moments**, by

$$\begin{aligned}\hat{\alpha}_1 &= \frac{m_3}{(m_2)^{3/2}} \\ \hat{\alpha}_2 &= \frac{m_4}{(m_2)^2}\end{aligned}\tag{4.8}$$

Our rationale for picking this “peculiar” notation, $\hat{\alpha}_1$, $\hat{\alpha}_2$, will be apparent in a few chapters. For the moment we need labels and $\hat{\alpha}_1$, $\hat{\alpha}_2$ are as good as any.

Consider any arbitrary change of origin and scale of any variable; that is, if x_i is the variable of interest, look at $y_i = a + (b \times x_i)$, for any values of a and b ,

$$\begin{aligned}
 \hat{\alpha}_1(y) &= \frac{m_3(y)}{(m_2(y))^{3/2}} \\
 &= \frac{b^3 m_3(x)}{(b^2 m_2(x))^{3/2}} \\
 &= \frac{m_3(x)}{(m_2(x))^{3/2}} \\
 &= \hat{\alpha}_1(x)
 \end{aligned}$$

$$\begin{aligned}
 \hat{\alpha}_2(y) &= \frac{m_4(y)}{(m_2(y))^2} \\
 &= \frac{b^4 m_4(x)}{(b^2 m_2(x))^2} \\
 &= \frac{m_4(x)}{(m_2(x))^2} \\
 &= \hat{\alpha}_2(x)
 \end{aligned}$$

These two new measures, $\hat{\alpha}_1$ and $\hat{\alpha}_2$, are invariant to changes in both scale and origin and so may correctly be called measures of shape. Before looking at some practical uses of our new tools, consider the following set of simple examples that will illustrate the ideas involved.

A key concept involved in any discussion of the higher moments is that, because a measure of the center of location and of spread, or scale, have already been determined, their confounding effects should be removed from the calculation of the higher moments. For example, the value of the third moment about the origin will reflect the effects of the degree of asymmetry, the center of location, and the spread. However, until the latter two effects are allowed for, you cannot distinguish the effect of asymmetry on the third moment.

We define four simple variables to illustrate. The four variables are x_a , x_b , x_c and x_d :

$$x_a = 1, 2, 3$$

$$x_b = 1, 2, 9$$

$$x_c = -3, 1, 2$$

$$x_d = -1, 0(10 \text{ times}), 1$$

Before beginning, sketch the frequencies as a line chart on any scrap of paper. We now calculate the four moments for each variable as well as the standardized third and fourth moments:

$$\begin{array}{ll}
 m'_1(x_a) = 2 & m_2(x_a) = \frac{2}{3} \\
 m_3(x_a) = 0 & m_4(x_a) = \frac{2}{3} \\
 \hat{\alpha}_1(x_a) = 0 & \hat{\alpha}_2(x_a) = \frac{3}{2} = 1.5
 \end{array}$$

$$\begin{aligned}
m'_1(x_b) &= 4 & m_2(x_b) &= \frac{38}{3} = 12.66 \\
m_3(x_b) &= 30 & m_4(x_b) &= \frac{722}{3} = 240.66 \\
\hat{\alpha}_1(x_b) &= \frac{30}{45.1} = 0.67 & \hat{\alpha}_2(x_b) &= \frac{240.66}{160.4} = 1.5 \\
\\
m'_1(x_c) &= 0 & m_2(x_c) &= \frac{14}{3} = 4.66 \\
m_3(x_c) &= -6 & m_4(x_c) &= \frac{98}{3} = 32.66 \\
\hat{\alpha}_1(x_c) &= \frac{-6}{10.1} = -0.59 & \hat{\alpha}_2(x_c) &= \frac{32.66}{21.78} = 1.5 \\
\\
m'_1(x_d) &= 0 & m_2(x_d) &= \frac{2}{12} = 0.166 \\
m_3(x_d) &= 0 & m_4(x_d) &= \frac{2}{12} = 0.166 \\
\hat{\alpha}_1(x_d) &= 0 & \hat{\alpha}_2(x_d) &= \frac{12}{2} = 6
\end{aligned}$$

The variables x_a and x_d are symmetric, so the third moment is zero, and we do not have to worry about the scaling problem. But is x_b five times more asymmetric than x_c as the raw third moment values would indicate? The standardized third moments are 0.67 and -0.59 , which seems to be much more reasonable in light of our drawings of the distributions. The unstandardized third moments are so different because the second moment of x_b is nearly three times greater than that of x_c ; that is, the unstandardized third moments have compounded the effects of the asymmetry with the differences in the values of the second moments, which indicate the degree of spread.

Looking at the fourth unstandardized moments, we would be misled into thinking that x_d has the smallest value for peakedness, that the value for peakedness for x_b is the largest by far, and that the value for x_c is much greater than that for x_d . All of these conclusions are wrong as we can see by examining our $\hat{\alpha}_2$ values, the values of the standardized fourth moments. Variable x_d has the largest value for peakedness as we might suspect if we look carefully at the relative frequency line chart. Interestingly, the values for peakedness for all the other variables are identical; again, we might suspect that fact from a glance at the frequency line charts.

Similarly, reconsider Figures 3.5 to 3.14 by examining the $\hat{\alpha}_1$ and $\hat{\alpha}_2$ values shown in Table 4.11 and the unstandardized moments in Table 4.9. Consider, for example, the $\hat{\alpha}_1$ values for Figure 3.11 and the lognormal in Figure 3.14; the latter is greater than the former, but for the unstandardized moments shown in Table 4.9 the opposite is true. This is due to the differences in the second moments. We noted previously that the lognormal distribution is a model for income distributions, and Figure 3.11 is the graph of the distribution for the film revenues. If these data are to be believed, we might wonder whether film revenues are less or more asymmetric than incomes.

Table 4.11 **List of Means, Second, and Standardized Third and Fourth Moments for Chapter 3 Histograms**

Figure	Subject	m_1	m_2	$\sqrt{m_2}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$
3.5	Die toss	4.80	6.00	2.50	-0.18	1.70
3.9	Final grades	67.30	132.00	11.50	-0.37	3.30
3.11	Film revenues (\$mm)	26.50	1050.50 (\$mm) ²	32.41	3.00	14.00
3.12	Gaussian	0.01	0.84	0.92	0.09	2.60
3.13	Weibull	0.97	0.02	0.13	-0.80	4.10
3.14	Beta (9,2)	0.83	0.01	0.11	-1.10	4.10
3.14	Arc sine	0.49	0.12	0.34	0.04	1.60
3.14	Uniform	0.52	0.09	0.30	0.25	1.80
3.14	Lognormal	1.81	6.66	2.60	5.32	40.30

Using the data in Table 4.4, we can calculate that the second moment of household income is 3.1 times 10^8 dollars, and the unstandardized third moment is 13.1 times 10^{12} dollars. However, the standardized third moment for household income is 2.4 and the standardized third moment for the film revenue data is 3.0, which is only a little larger.

Recall that the household incomes data in Table 4.4 involve some considerable approximation, especially in that the very highest incomes, although of very low relative frequency, were not recorded in the table. We expect that the actual standardized third moment is greater than that calculated. In any event, we can conclude that film revenues are not substantially more skewed than incomes generally.

What of peakedness for these data? The fourth moment for the household income data is 114.2 times 10^{16} dollars, but the standardized fourth moment is 11.9. The corresponding values for the standardized fourth moments shown in Table 4.11 are 13.9 for the film revenue and 40.3 for the lognormal distribution. Recalling once again that the nature of our approximations for the household income data is likely to underestimate the fourth moment, we can put the film revenue results into some perspective; they are at least not substantially more peaked than the household income data.

Three distributions in Table 4.11 seem to have similar values for the standardized fourth moments. Figure 3.5 was for the die-tossing experiment, and we would expect intuitively that the distribution of outcomes would be flat and not peaked. The uniform distribution is the ultimate in flat distributions. The arc sine distribution is U-shaped; such a distribution might be thought of as anti-peaked, so it should have a very low value for the standardized fourth moment. The $\hat{\alpha}_2$ values are, respectively, 1.7, 1.8, and 1.6. The Weibull and beta distributions shown in Figures 3.13 and 3.14, respectively, have about the same standardized fourth moments as we would expect from looking at the figures; both have $\hat{\alpha}_2$ values of 4.1.

Some Practical Uses for Higher Moments

At last we are ready to compare our film revenue data to see which film studios we want to back. Before looking at the listing of all four moments for the nine different film studios shown in Table 4.12, it would be helpful for you to refresh your memory